Apr 9th, 12:00 AM

NSF Data Management and Public Access Initiatives

Anne Maglia
National Science Foundation

NSF Data Management and Public Access Initiatives

e-Science Symposium
9 April 2015

Anne Maglia
Division of Biological Infrastructure
National Science Foundation
amaglia@nsf.gov

DMP → Public Access
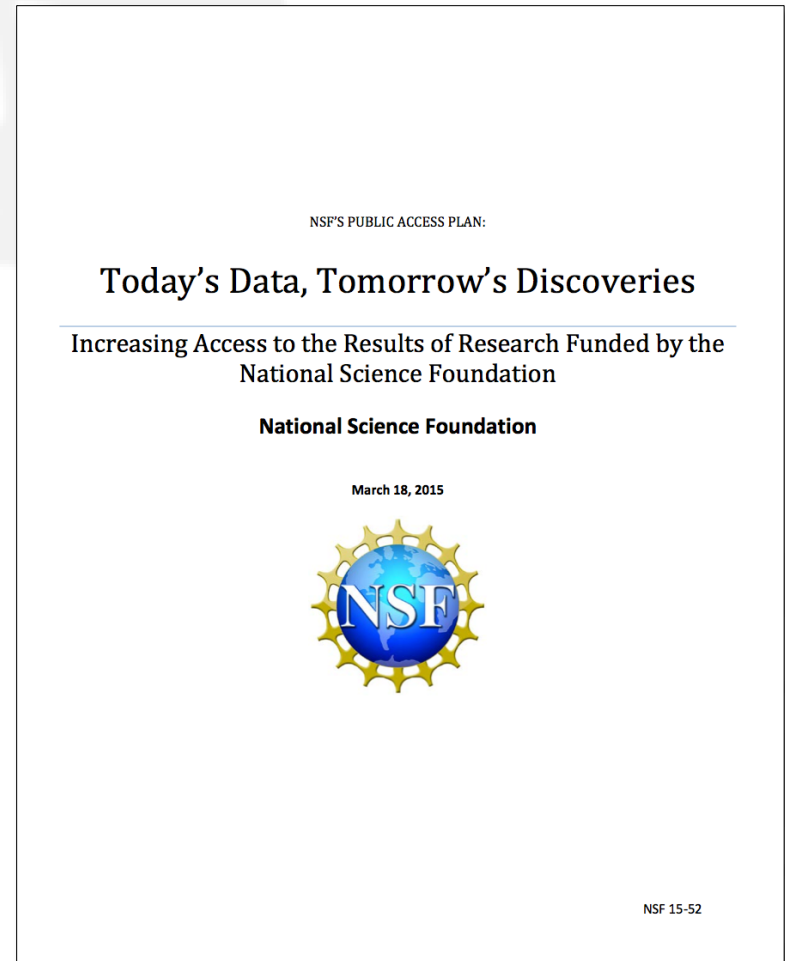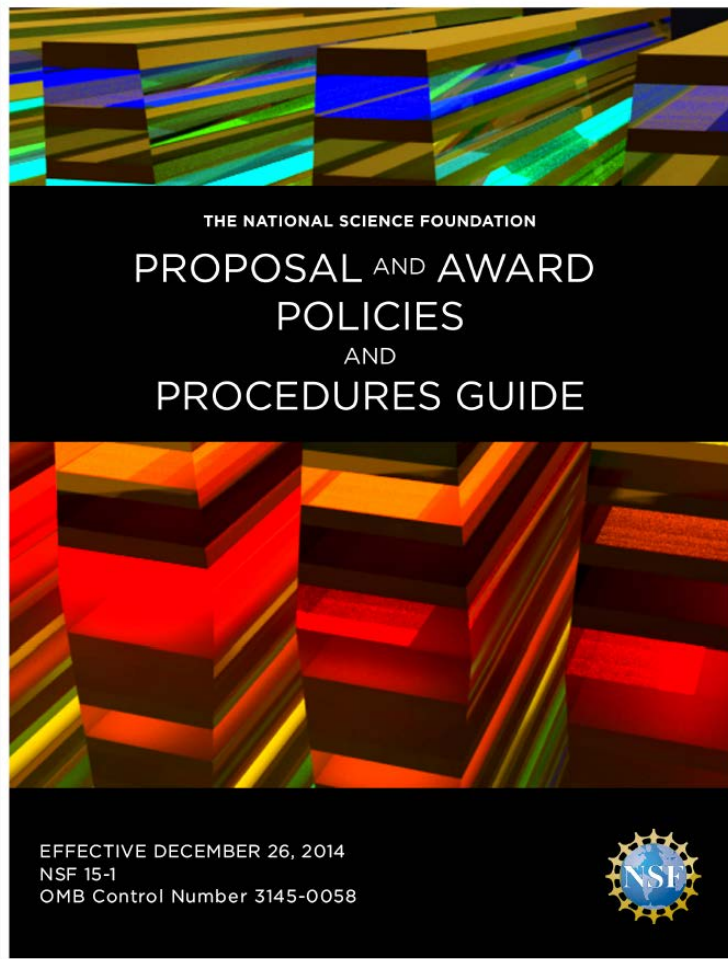
Outline

- Policy:
  - NSF Data Management Plan
  - NSF/BIO's guidance for DMP and post-award reporting
- NSF/BIO's Strategy:
  - Award Management
  - Assessment
  - Joint Visioning with CISE/ACCI
  - Investment
    - Planning and Community Engagement
    - Implementation
- Policy:
  - NSF Public Access Plan

NSF Data Sharing Policy

"Investigators are expected to share with other researchers, at no more than incremental cost and within a reasonable time, the primary data, samples, physical collections and other supporting materials created or gathered in the course of work under NSF grants. Grantees are expected to encourage and facilitate such sharing. "

NSF Data Management Plan
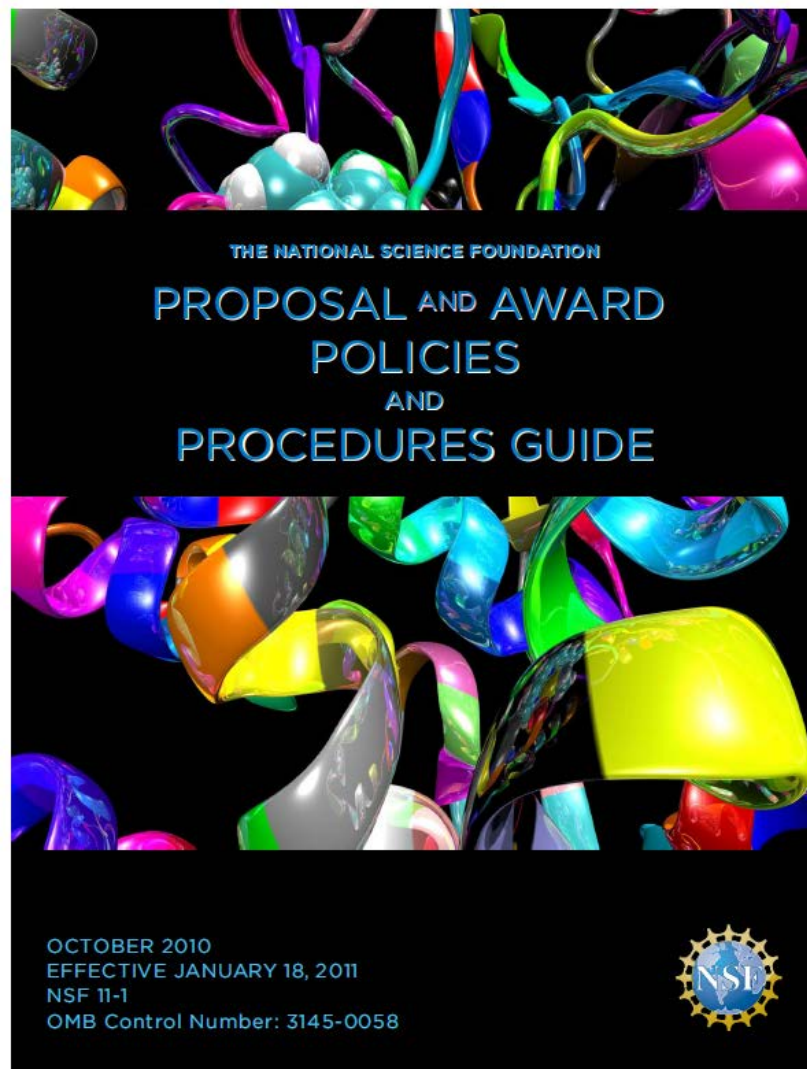
NSF Data Management Plan

- 2 supplemental pages
- FastLane requirement/Return Without Review
- Original data, metadata, and software/code
- Details on how data will be managed and disseminated, past performance
- Sufficient for replication

NSF/BIO DMP Guidance

- Sub-disciplines have their own standards
- Accepted norms for timely distribution vary
- Follow best practices in proposed sub-discipline
- Address how standards, best practices, and timeliness will be met

http://www.nsf.gov/bio/pubs/BIODMP061511.pdf

Bio Directorate for Biological Sciences

UPDATED Information about the Data Management Plan Required for all Proposals (2/20/13)

BACKGROUND

The National Science Foundation (NSF) required a data management plan (DMP) for all full proposals submitted, or due, to NSF on or after January 18, 2011. For the full policy implementation, see the NSF Grant Proposal Guide (GPG) Chapter II.C.2.j, in the NSF Proposal and Award Policies and Procedure Guide.

REQUIREMENT

Proposals must include a supplementary document of no more than two pages labeled "Data Management Plan". Any specific instructions and exceptions to the two page limit will be found in specific Program Solicitations. The DMP is NOT part of the 15 page Project Description. Even if no data will be produced (e.g., a workshop proposal), a DMP should be submitted that states: "No data are expected to be produced from this project."

NSF/BIO DMP Guidance

1. Describe data collected and metadata formats and standards
2. Describe data storage and preservation resources and facilities (incl. post award)
3. Describe data/metadata dissemination methods (incl. post award)
4. Describe data sharing and public access policies for (incl. privacy, confidentiality, security, IP, etc.)
5. Describe roles and responsibilities re: data management (incl. contingency and post award)

NSF/BIO Post-Award Reporting Guidance

- Annually: data management/sharing progress
  - Identifiers, accession numbers, citations, etc.
- Final report: implementation of the DMP
  - data produced
  - data retained post-award
  - dissemination and verification of sharing
  - format in which data and metadata are available
  - location of data for long-term public access
- Future proposals > "Results of prior NSF support"

Outline

- Policy:
  - NSF Data Management Plan
  - BIO's guidance and post-award management
- NSF/BIO's Strategy:
  - Award Management
  - Assessment
  - Joint Visioning with CISE/ACCI
  - Investment
    - Planning and Community Engagement
    - Implementation
- Policy:
  - NSF Public Access Plan

Award Management

- Solicitation/award letter language and conditions
- ABI Sustaining track
- Cooperative agreements

Advances in Biological Informatics (ABI)

PROGRAM SOLICITATION
NSF 12-567

REPLACES DOCUMENT(S):
NSF 10-567

National Science Foundation
Directorate for Biological Sciences
Division of Biological Infrastructure

Full Proposal Deadline(s) (due by 5 p.m. proposer's local time):

September 10, 2012

August 13, 2013

Second Tuesday in August, Annually Thereafter

IMPORTANT INFORMATION AND REVISION NOTES

Revision Notes:

1. Program scope and description changed to create total of three classes of awards: a new class called Sustaining awards, in addition to the existing classes of Innovation awards and Development awards.

2. Language pertaining to available funds and anticipated numbers of awards has been changed to reflect the approximate balance between numbers of Innovation, Development, and Sustaining awards.

3. Proposal Preparation and Submission Instructions have been edited to accommodate changes related to the three submission types

4. Added additional merit review considerations for Sustaining awards.

5. Changes to availability of funds for workshops, conference support, and REU supplements.

Assessment

- Internal analysis of BIO DMPs (Cragin, Cak)
  - assessed DMPs for 5 Guidance items
- PD surveys
- Sample DMPs for program requirements (LTER)

Preserving BIO-Funded Research Data:
Leveraging Research Data Management and the Public Access Policy

Joint Visioning

- BIO AC & CISE ACCI: April 23rd
- Discuss major BIO data infrastructure investments
- Identify priorities, gaps, opportunities

NSF Advisory Committee for Cyberinfrastructure

October 29, 2014 - October 30, 2014
Advisory Committee for Cyberinfrastructure
Advisory Committee Meeting

April 2, 2014 - April 3, 2014
Advisory Committee for Cyber
Advisory Committee Meeting

Directorate for Biological Sciences Advisory Committee (BIO AC)

ABOUT THE BIO Advisory Committee

The Biological Sciences Advisory Committee (BIO AC) advises the Directorate (BIO) on such issues as:

- How BIO's mission, programs, and goals can best serve the scientific community;
- Institutional administration and policy;
- How BIO can promote quality graduate and undergraduate education in the biological sciences;
- Priority investment areas in biological research; and
- Government Performance and Results Act, including Committees of Visitors.

Sample Investments

- <u>GoLife</u>: integrated data and information resources for comparative biology

- <u>Sustainability Training</u>: Ecological Society of America peer-based course for sustainable biological infrastructure (1340550)

- <u>GBIF, IPBES</u>: U.S. participation in the Global Biodiversity Information Facility and the Intergovernmental Platform for Biodiversity and Ecosystem Services (1451069,1415669)

Public Access Workshops

- Addressing changing practices around the publication of biological data (1449499)

- Reducing barriers for the management, integration, and public sharing of large and complex data among biologists (1450894)

- Guidelines on preparing animal behavior data for sharing, publication, and access (1451110)

- Management and preservation of audio and video data (1451374)

iDigBio

Integrated Digitized Biocollections

About iDigBio | Research | Technical Information | Education

Google™ Custom Search [Search]

Log In | Sign Up

Making data and images of millions of biological specimens available on the web

27,187,981 Specimen Records

4,317,649 Media Records

391 Recordsets

[Search the Portal]

Why digitization matters
More about what we do and why

Digitization
Learn, share and develop best practices

Sharing Collections
Documentation on data ingestion

Working Groups
Join in, contribute, be part of the community

Proposals
New tool and workshop ideas

Citizen Scientists
How can you help biological collections?

Researchers
Learn about research directions

Collections Staff
Learn how your collection can benefit from our work

Teachers & Students
Download lesson plans about using digitized specimens

Upcoming Events

Citizen Science course @FSU
04-08-2015

Webinar: Practitioner tools and resources for evaluating learning outcome in citizen science
04-08-2015

iDigBio @ The National Association for Research in Science Teaching Annual Meeting
04-11-2015 to 04-14-2015

International Digitization Summit
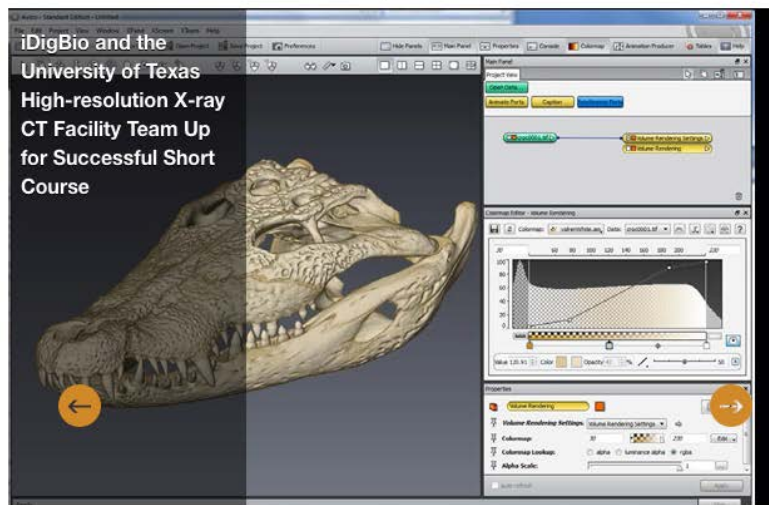04-13-2015 to 04-17-2015

more events >>

News

iDigBio Listed as a Resource on the Linking Florida's Natural Heritage Page
Post date: 03-20-2015

New Position with NSF: Division Director, Division of Biological Infrastructure
Post date: 03-11-2015

iDigBio and the University of Texas High-resolution X-ray CT Facility Team Up for Successful Short Course

www.iplantcollaborative.org

iPlant Collaborative™

Platforms & Tools | News | Events | Learning Center | About | Help | Status | Contact Us

Search

iPlant: from data to discovery

The iPlant Collaborative is where scientists in all domains of life sciences can connect to public datasets, manage and store their own data and experiments, access high-performance computing, and share results with colleagues.

Take a Tour

Platforms and Tools
See All

Discovery Environment
Use hundreds of bioinformatics Apps and manage data in a simple web interface

DNA Subway
Take DNA Subway to teach classroom-friendly bioinformatics for genome analysis, DNA Barcoding, and RNA-Seq

Bisque Image Analysis Environment
Exchange, explore, and analyze biological images and their metadata

Atmosphere
Create a custom cloud-based scientific analysis platform or use a ready-made one for your area of scientific interest

Data Store
Store, manage, access, and share all the data related to your research

User Portal

About iPlant

The Project
Building cyberinfrastructure for life sciences research

Science
Enabling discovery in the age of Big Data

Success Stories
How users and projects have used iPlant

Publications
Recent publications and how to cite iPlant

Help
Find answers to common questions

News

iPlant Emergency Maintenance: Tuesday, April 7, 8AM - 5PM MST
Discovery Environment and Atmosphere services may...

[UPDATED - COMPLETE] iPlant Maintenance: Tuesday, March 31, 8AM - 5PM MST
Discovery Environment version 1.9.6 has been...

Supercomputers, Statistical Software, and Science: Bridging the Gap to Study Plant Adaptation
Scientists took advantage of the iPlant...

A Tug-of-War Between Plants and Pathogens
Analyses completed using the iPlant Collaborative...

MORE...

Upcoming Events

iPlant Tools & Services Workshop - University of Georgia
University of Georgia, Athens, GA

iPlant Tools & Services Workshop - Santa Fe Community College
Santa Fe, NM

Genomics in Education Workshop -Santa Fe Community College, Santa Fe, NM
Santa Fe, NM

Learning Center

Get Started
First things to do when you get your account

Platform and Science Tutorials
Step-by-step guides for our platforms and specific analyses

Ask iPlant
Post your science and support questions

Powered by iPlant

Galaxy

Outline

- Policy:
  - NSF Data Management Plan
  - BIO's guidance and post-award management
- NSF/BIO's Strategy:
  - Award Management
  - Assessment
  - Joint Visioning with CISE/ACCI
  - Investment
    - Planning and Community Engagement
    - Implementation
- Policy:
  - NSF Public Access Plan

NSF Public Access Plan

https://www.nsf.gov/publications/pub_summ.jsp?ods_key=nsf15052

Requirements/Constraints

- Provide public access to results of NSF awards
- Minimize burden on PIs, NSF
- Embed public access into existing workflows
- Leverage existing systems
    - e.g., e-Jacket; research.gov, external services
- Extend to multiple products of NSF-funded research
- Minimize cost

NSF Public Access Plan

- As of Jan 2016, requires publications from awards to be made publicly available within one year

- (Waivers to embargo possible)

- DMP requirements, allowance for costs, data citation stay same

- Calls for community engagement in use of identifiers, metadata, standards

NSF Public Access Plan

- Leverages current programs, policies, and systems (incl. DoE Pages)

- Establishes oversight Working Group

- Establishes a website for Plan, feedback, and FAQs

- Calls for regular updates to NSB and OSTP/OMB

- Allows for expansion for other products of NSF-funded research

Investigator/author

1. NSF Credentials
2. NSF Award ID
3. DOI (publisher)
4. Copy of paper (PDF/A)

Research.gov

NSF credentials

Research.gov submit

Research.gov report

Select Award

PI enters DOI (publisher supplied)

NSF Awards

Auto-populates the report

Cross Ref

Auto-populates the metadata

PI links award to record & uploads AM

Author completes report & submits

Ingested into NSF Public Access Repository

Visible in e-Jacket

Public: Display, read, download, analyze

Program Officer

Questions?

amaglia@nsf.gov
703-292-7380
@ammaglia